

Can I trust artificial intelligence with questions about my health?

INVESTIGATING THE ACCURACY OF AI-GENERATED RESPONSES TO SIMPLIFIED OR DEMOGRAPHIC-SPECIFIC MEDICAL QUESTIONS

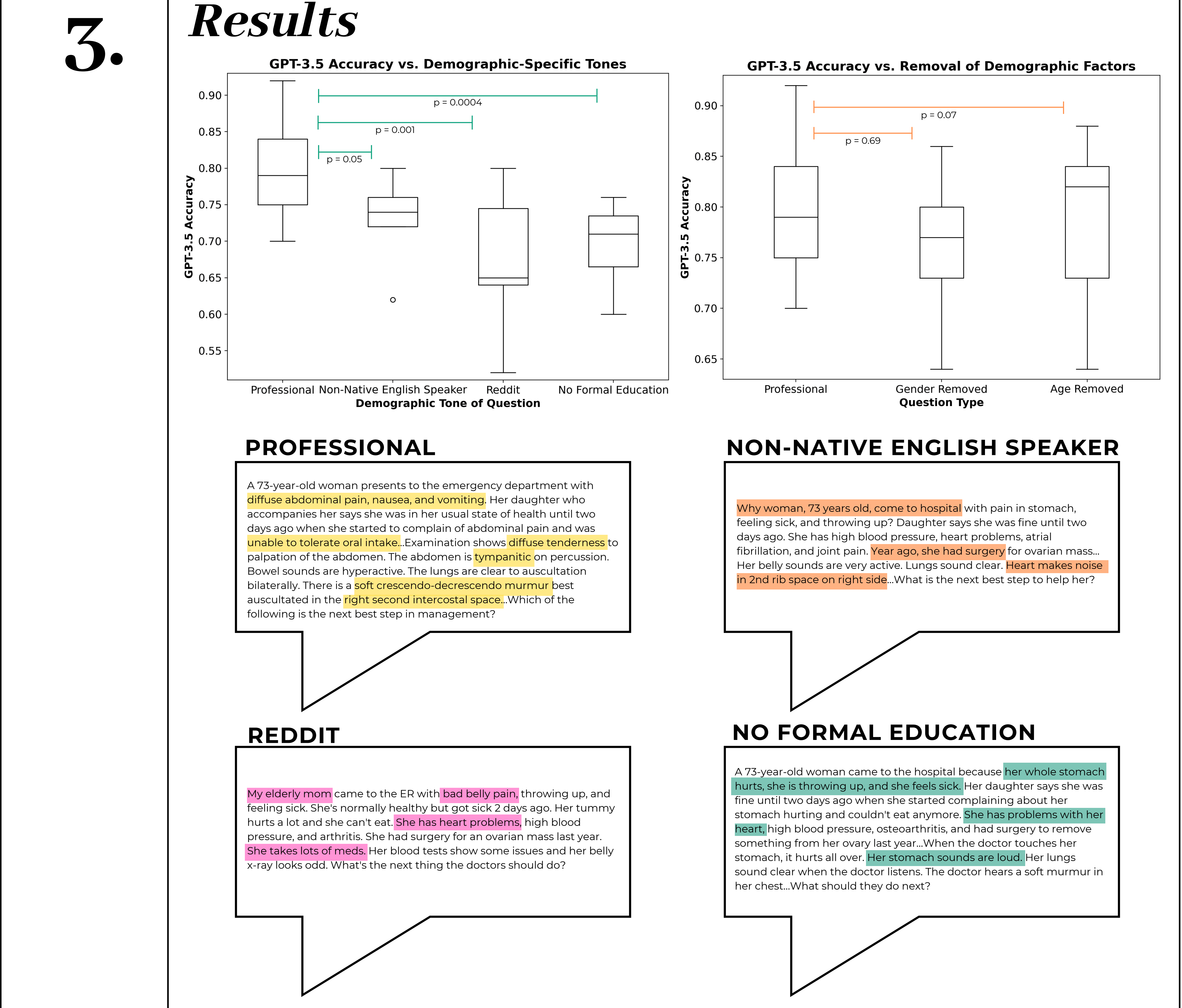
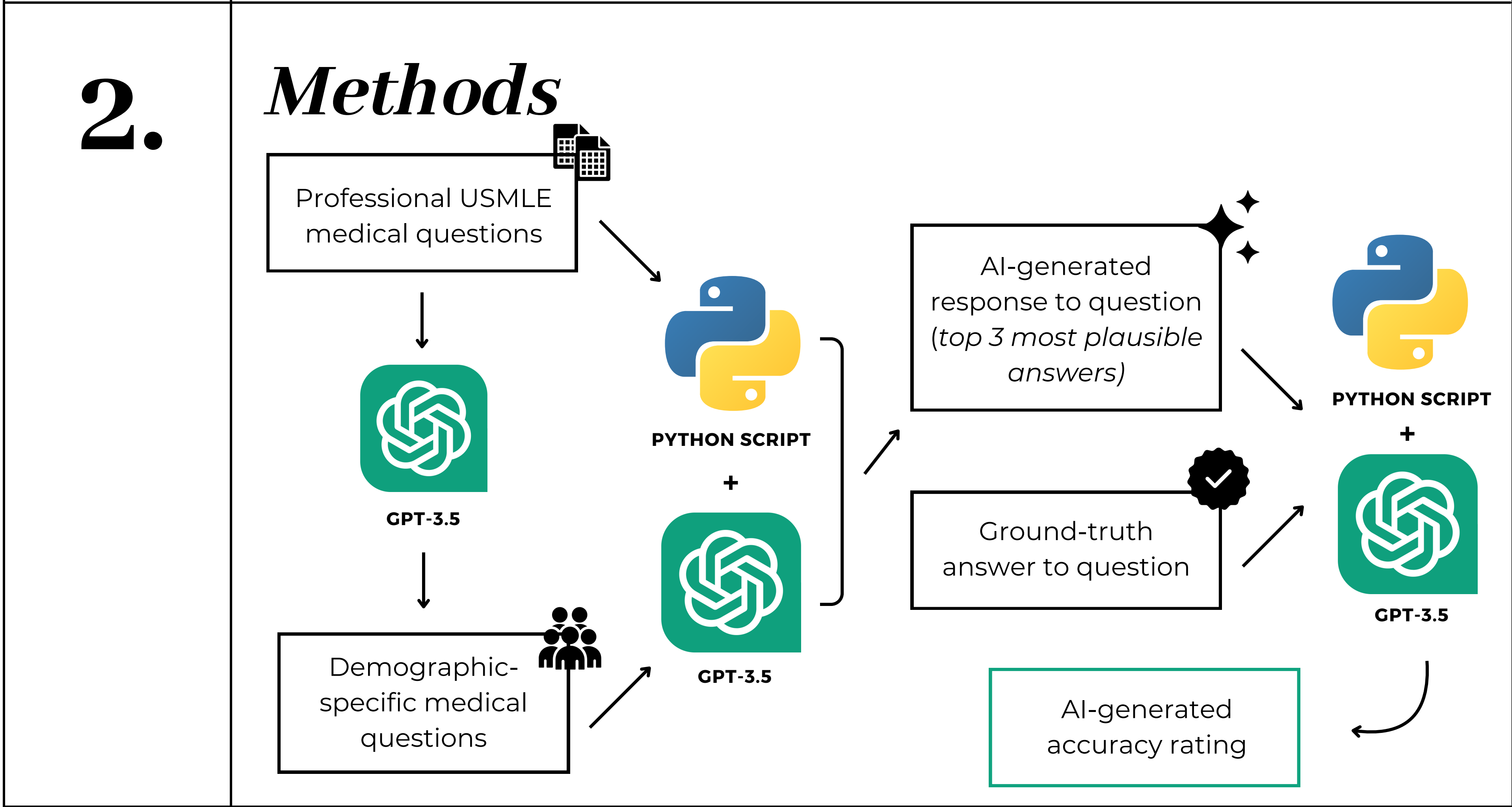
1.

Motivation

The performance of GPT-3.5, a large language model (LLM), has been tested extensively on professional medical questions from the United States Medical Licensing Exam (USMLE) and other similar datasets. However, GPT-3.5 has not been tested on simplified or unprofessional medical questions that are likely to come from everyday patients turning to AI platforms for advice about their health.

?

- Will the widely recognized performance of GPT-3.5 on medical questions decrease if questions are simplified or asked in unprofessional, realistic tones?
- If so, which factors make significant contributions to the decrease in performance?



4.

Discussion

While removing age and gender from professional medical questions did not impact GPT-3.5's accuracy, rephrasing the questions in the tone of a non-Native English speaker, Reddit user, or someone with no formal education resulted in a decrease in the LLM's accuracy. Our results suggest that GPT-3.5 is not yet trained to appropriately respond to medical questions asked by everyday patients.

5.

Future work

- Include multiple choice answers in the prompt to restrict the agency of GPT-3.5's response
- Seek out tone-specific examples to train GPT-3.5 on like we did with Reddit data
- Rerun the experiment with GPT-4, a newer and more accurate version of the LLM

6.

References

Matzinger T, Ritt N, Fitch WT (2020) Non-native speaker pause patterns closely correspond to those of native speakers at different speech rates. PLOS ONE 15(4): e0230710. <https://doi.org/10.1371/journal.pone.0230710>

Nori, H., King, N., McKinney, S. M., Carignan, D., & Horvitz, E. (2023). Capabilities of GPT-4 on Medical Challenge Problems. arXiv preprint arXiv:2303.13375.

Pereira, A., Ortiz, K.Z. Language skills differences between adults without formal education and low formal education. Psicol. Refl. Crit. 35, 4 (2022). <https://doi.org/10.1186/s4155-021-00205-9>

<https://thenounproject.com/>